1972

Establishing criteria in programmed learning.

Madson, Gerald Gerhart.

Monterey, California. Naval Postgraduate School

ESTABLISHING CRITERIA IN PROGRAMMED LEARNING

Gerald Gerhart Madson

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

ESTABLISHING CRITERIA IN PROGRAMMED LEARNING

by

Gerald Gerhart Madson

Thesis Advisor: J. K. Arima

September 1972

Establishing Criteria in Programmed Learning

by

Gerald Gerhart Madson
Captain, United States Marine Corps
B.S., Marquette University, 1963

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1972

ABSTRACT

Recent years have seen many attempts to program learning based on many principles ranging from intuition to mathematics. The area of criterion establishment for advancement or conversely additional training at the same or lower levels remains primarily in the realm of intuition. The intuition used is only as good as the experience and background of the program designer however. In the case of a very experienced program designer the criteria may be very efficient. In the case of less experienced program designers the criteria are usually arbitrary. This work describes a method of analyzing learning programs and determining mathematically sound criteria. The mathematical foundation for this analysis is the Markovian learning model as opposed to the linear learning model.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

A. GENERAL

This work was stimulated through the author's association with the Behavioral Sciences Institute of Monterey, California. Members of the staff of the Institute have developed a procedure for preparing language training programs [Gray and Ryan 1973]. This procedure "Language Training Through Programmed Conditioning" is based on a mathematical principle or structure. The total language program consists of ninety sub-programs covering different language deficiencies which are the result of various causes such as deafness, mental retardation, non-English speaking, etc. Each sub-program consists of a series of steps that sequentially train a student until he reaches a desired objective. The total program is designed to test the student in a continuous manner to determine what sub-programs or parts (steps) he needs, when he has completed any step or complete sub-program, and when he is having excessive problems that require stepping back (branching) or other forms of special help.

The procedure, though developed for language training, probably has wide application. These applications will appear as more people, in varied fields, become familiar with the procedure and investigate its appropriateness to their specific areas of interest. Along this vein it is also hoped that the method developed herein, of analyzing the program and

determining criteria to be used in the administration of the program, will have wider application.

B. PROGRAMMING

The objective of an intellectually oriented training program is to move a subject from a specific unlearned state (U) to a learned state (L). The process used to achieve this goal is Stimulus-Response-Consequence, normally referred to as operant conditioning. In most situations the desired learned state (L) can be achieved most effectively by reducing the unlearned state (U) to a number of small steps (states) which are learned in a sequence and build up to the ultimate goal. This latter process is called programming. Throughout this work the term presolution will be used interchangeably with the unlearned state and the term solution with the learned state.

C. CRITERIA

The subject of criteria arises when the programmer desires to measure the success of a given program in moving a specific subject from the unlearned to the learned state. Basically the question is: "What measure should be used for determining when a subject has achieved the learned state (L)?" A second aspect of criteria arises when it appears that a subject is having an excessive amount of trouble with a particular step. In such a case it may be desirable to alter the program for the subject, possibly reducing the step to several less difficult substeps. The first criterion is referred to as advancement criterion and the second is referred to as

branching criterion. A difficulty in establishing these criteria results from the generally accepted idea that there is no exact (deterministic) method of establishing criteria. Alternately learning is viewed as being probabilistic (stochastic) in nature and therefore the criteria must be established within the realm of probability theory. To further complicate the situation there is not a universally accepted learning theory and therefore the criteria, if determined statistically, will be a function of the learning theory the programmer subscribes to.

Some possible measures for criteria are: total number of correct responses, total number of incorrect responses, total number of trials, plus the associated ratios and percentages. A trial is one Stimulus-Response-Consequence sequence. Of these it seems intuitively appealing to select total number or percentage of correct responses as an indication of when to advance a subject. More specifically it is common practice to use a string of consecutive correct responses for advancement. Determining the number in this string and the confidence in that number is required in establishing the first criterion. Determining a maximum acceptable number of trials and/or errors per step, and the associated confidences, is the procedure used to establish the second criterion.

II. MODEL DEVELOPMENT

A. BACKGROUND

In a brief historical review of mathematical learning theory, Atkinson, Bower, and Crothers (1965, pp. 19-24), indicate that the treatment of learning using probabilistic models began with Thurstone in 1919. From 1919 through 1950 there were numerous probabilistic models proposed and tested, all related to various specific learning situations. Since 1950 there has been a great deal of work done in the area of stochastic learning, and two theories have evolved as the primary contenders for acceptance: the linear model and the Markov model. Basically the linear model states that a subject has a probability of success on any trial n which is given by the following equation:

$$P_n = 1 - (1-P_1)(1-\theta)^{n-1} \qquad (1)$$

where P_1 is his initial probability of success and θ is his learning rate. The Markov model takes a quite different approach. Basically the theory says that if a subject is in an unlearned state (U) then the probability of a correct response is g (guess), and if the subject is in the learned state (L), then the probability of a correct response is 1. The probability of going from the unlearned state to a learned state on any presolution trial is usually denoted c. The probability of a correct response on any trial n is usually given by:

$$P_n = 1 - (1-g)(1-c)^{n-1} \tag{2}$$

Comparison of Eqs. (1) and (2) indicates that their forms are exactly the same. The difference in these two equations lies in their theoretical background and the meaning of the parameters. Equation (1) states that a subject starts with a probability P_1 of making a correct response on the first trial. The probability of success on the second trial is greater, due to the incremental learning achieved on the first trial. This linear process continues indefinitely and the subject's probability of success asymptotically approaches 1. Equation (2) states that on each presolution trial a subject has a probability c of going into solution. Once in solution the subject will remain in solution and always respond correctly (P_n=1). While in presolution the subject has a probability g (guess) of responding correctly and this probability does not change. The form of these two equations and other interesting comparisons are made in "Learning: Gradual or All-or-None," [Restle and Greeno 1970, Chap. 2]. Based on their analysis Restle and Greeno stated, " . . . the all-or-none theory is the most interesting, and we think it is the one most deserving of future work," (p. 79).

Pilot research involving a computer simulation of the linear model yielded a set of criteria for use in the language training programs of the Institute. Informal observation of the data on which the linear model was based suggested that the linear model might not be appropriate. Further

study [1] of data from students trained on these and other programs showed a definite tendency toward the Markov principles of stationarity and independence [Atkinson, and others, 1965, pp. 39-49; Coombs, Dawes, and Tversky 1970, pp. 296-297]. Based on these results the linear model was set aside and this work started based on the Markovian (All-or-None) principle.

B. ASSUMPTIONS

The following assumptions are necessary to the development:

1. The learning process involved in the programs developed by the Institute are Markovian in nature.

2. The subject can be correct on the first trial of any step by either: (1) being in solution prior to the trial, (2) going into solution because of the information presented in the first stimulus, or (3) guessing correctly in presolution. This assumption modifies Eq. (2) in that Eq. (2) contains the restriction that for the subject to be correct on the first response, he must guess correctly, therefore not allowing the possibility of being in solution (the learned state) on the first trial. Allowing for the possibility that the subject is in solution on the first trial [Atkinson, and others, 1965, footnote p. 55], appears to be a more realistic approach and was used in this work.

---

[1] Gray, B. B.; Carrier, J. K., Jr.; Bradfield, R. H.; and Rigg, L., The Discrete Effects of Response Contingent Reinforcement During Skill Acquisition, submitted to Journal of Applied Behavior Analysis for publication.

3. The g factor (guess) in presolution is a function of the conditioning sequence (step) and the subject.

4. The c factor (probability of going into solution) is a function of the step and the subject.

5. g and c are constant over any step for a given subject.

6. The set of outcomes form a homogeneous Markov chain.

$$\begin{array}{ccc} & L_{n+1} & U_{n+1} \\ L_n & 1 & 0 \\ U_n & c & (1-c) \end{array} \cdot \begin{array}{c} P_r(\text{Correct}|\text{Row State}) \\ \begin{bmatrix} 1 \\ g \end{bmatrix} \end{array} \qquad (3)$$

$$n = 0,1,2, \ldots$$

$$g \leq 1, \; c \leq 1$$

C. MODEL

The equations developed in this work are a combination of what is presented in Refs. 1 - 3 and some original work. Since there is so much overlap it is not possible to give credit to any one source, instead the equations will be given with explanations and generally acknowledging the credit to all. The first development is the probability of errorless response given that the subject is known to be in the unlearned state (U). This state is assumed on the first trial and known to exist if an error occurs. If a subject is put on a program step and no error occurs before reaching the advancement criterion, then there is no way of determining whether he was in solution (learned state, L) initially, or was in presolution (unlearned state, U) initially and performed as follows:

$$P[\text{Errorless}] = c + g(1-c)c + g^2(1-c)^2c + \ldots$$

$$\therefore \lim_{n\to\infty} P[\text{Errorless}] = \frac{c}{1-g(1-c)} \tag{4}$$

For future reference, this probability will be called ρ (rho), the probability of errorless response given that the subject is in presolution. (Being in presolution is demonstrated anytime a subject makes an error.) The above development says that the subject either goes into solution on the first trial; stays in presolution, guesses correctly and then goes into solution; stays in presolution twice, guesses correctly twice and then goes into solution; etc. The development indicates that the subject eventually goes into solution if errorless response is achieved after an error. The reader familiar with the Markov theory will note that the term relating to remaining in presolution and having errorless response was omitted in developing Eq. (4). The omission was committed since the term $g^n(1-c)^n$ goes to zero in the limit as n approaches infinity.

The next development will be the expected number of errors given g and c. The development relies on the parameter ρ. The probability that the total number of errors is exactly k equals:

$$P[E = k] = \sum_{i=0}^{\infty} \binom{k+i}{i} g^i(1-c)^i(1-g)^k(1-c)^k c$$

If this abbreviated mathematical statement were written out, it would represent every feasible combination of events in which exactly k errors can occur.

$$\therefore\ P[E = k] = (1-g)^k (1-c)^k c \sum_{i=0}^{\infty} \binom{k+i}{i} g^i (1-c)^i$$

Referring to a standard math table e.g. CRC [13th Edition, p. 463]. This equation then reduces to

$$P[E = k] = [(1-g)(1-c)]^k\ c\ [\frac{1}{1-g(1-c)}]^{k+1}$$

$$= [\frac{(1-g)(1-c)}{1-g(1-c)}]^k\ [\frac{c}{1-g(1-c)}]$$

$$= [1 - \frac{c}{1-g(1-c)}]^k\ [\frac{c}{1-g(1-c)}]$$

Referring to Eq. (4) and substituting

$$\therefore P[E = k] = (1-\rho)^k \rho \tag{5}$$

In words Eq. (5) says that there were k possible errorless response strings broken by the occurrence of an error and then there were no more errors. Since the probability of an errorless response string is ρ, given that the subject is in presolution, it follows that not errorless response is equal to $(1-\rho)$. This takes into account all possible numbers of correct responses before the error response that breaks the string. The occurrence of an error demonstrates presolution and also allows another possible string of errorless response, which is independent of the length or number of previous strings and is only dependent on the knowledge of being in presolution, i.e., the occurrence of an error. In summary then Eq. (5) says that the probability of exactly k errors occurring is equal to not errorless response k times and then errorless response.

$$E[E] = \sum_{k=0}^{\infty} k \cdot P[E=k] = \sum_{k=0}^{\infty} k \cdot \rho(1-\rho)^{k}$$

$$= \rho(1-\rho) \sum_{k=0}^{\infty} k(1-\rho)^{k-1}$$

$$= \frac{\rho(1-\rho)}{\rho^{2}} = \frac{(1-\rho)}{\rho} \qquad (6)$$

$$\therefore E[E] = \frac{(1-g)(1-c)}{c} \qquad (7)$$

The next development is the expected trial number of the last error. The probability that the last error occurred on trial t equals:

$$P[T=0] = \rho$$

$$P[T=t] = (1-c)^{t}(1-g)\rho \qquad (8)$$

$$t = 1,2,3,\cdots$$

In words, Eq. (8) says that there were t presolution trials indicated by an error on trial t and then errorless response. This probability statement allows for any sequence (order) or number of correct and incorrect responses up to trial t. The only required knowledge is that an error occurs on trial t and then no more errors. It should be noted that the probabilities given by Eqs. (5) and (8) are both decreasing functions. This corresponds with what will be observed about Figs. 2 and 1 respectively.

$$E[T] = \sum_{t=0}^{\infty} t \cdot P[T=t]$$

$$= 0.\rho + \sum_{t=1}^{\infty} t \cdot P[T=t]$$

$$= \sum_{t=1}^{\infty} t \cdot (1-c)^{t}(1-g)\rho$$

$$E[T] = \rho(1-g)(1-c) \sum_{t=1}^{\infty} t \cdot (1-c)^{t-1}$$

$$= \frac{\rho(1-g)(1-c)}{c^2}$$

$$= \frac{\rho}{c} \cdot E[E] = \frac{E[E]}{1-g(1-c)} \tag{9}$$

$$\therefore E[T] = \frac{(1-g)(1-c)}{[1-g(1-c)]c} \tag{10}$$

With Eqs. (6) and (9) it only requires a little manipulation to get estimates of g and c:

from Eq. (6)

$$\rho = (E[E]+1)^{-1}$$

from Eq. (9)

$$c = \rho \cdot \frac{E[E]}{E[T]}$$

$$\therefore c = \frac{E[E]}{E[T](E[E]+1)}$$

$$\therefore \hat{c} = \frac{E}{T(E+1)} \tag{11}$$

In Eq. (11) the expected number of errors E(E) is replaced by the observed number of errors and similarly the trial of the last error T replaces E(T). This might seem out of ord[illegible] to

some, in that a random variable is substituted for its expected value (mean or first moment). Some thought, however, will point out that the observed values of E and T are the mean when the sample size is one. Each subject/step combination represents a unique data point that cannot be repeated since the completion signifies the transition from the unlearned state (presolution) to the learned state (solution), and g and c factors are no longer applicable when the subject is in solution. This procedure represents a special case of the application of method of moments in estimating the parameters of a distribution:

$$\text{from Eq. (9)} \qquad g = \frac{1 - \frac{E[E]}{E[T]}}{(1-c)}$$

Substituting E=E(E) and T=E(T) and utilizing Eq. (11):

$$\hat{g} = \frac{\frac{T - E}{T}}{1 - \frac{E}{T(E+1)}}$$

$$= \frac{(T-E)(E+1)}{T(E+1)-E}$$

$$\therefore \hat{g} = 1 - \frac{E^2}{T(E+1)-E} \qquad (12)$$

If E is much greater than 1 (E>>1) then

$$\hat{c} \approx \frac{1}{T}$$

Also

$$\hat{g} = 1 - \frac{E}{(T-1) + \frac{T}{E}}$$

Since T is greater than or equal to E($T \geq E$)

$$\hat{g} \geq 1 - \frac{E}{T}$$

These equations state that c is approximated by the reciprocal of the number of the trial of the last error. This is intuitively appealing as it states the larger the c factor (probability of going into solution), the fewer expected number of trials. The estimate of g is slightly larger than the apparent presolution accuracy. This comes from the concept which allows the possibility of presolution correct responses after the last error. If the number of these responses were known, the estimated presolution accuracy rate ($\hat{g}$) would increase above what can be estimated using only the trials to the last error. The required correction factor is taken into account in Eq. (12).

PREFACE — SECTION III

During the final proofreading of this thesis it was discovered that the computation of $\hat{g}$ (Appendix B and Figs. 4, 7, and 8) had been done using: $\hat{g} = 1 - \frac{E^2}{T(E+1)+E}$, instead of $\hat{g} = 1 - \frac{E^2}{T(E+1)-E}$ (Eq. 12).

The second equation is the valid equation as it appears in the text. Specific reference to $\hat{g}$, as it applies to the data analyzed, must therefore be considered in light of the first equation, e.g., pages 20 and 21.

The effect of using the first equation is to cause an across-the-board increase in $\hat{g}$. The general conclusions are still valid but the specific conclusions as regards the data analyzed, must be viewed in light of this new information. Examples of the effect of this change are: (1) a general shift of Figs. 4, 7, and 8 to the left, (2) a general decrease in the mean values of $\hat{g}$ listed in Tables I and II, (3) a general decrease in the estimated values listed in Table III. This is based on observing that the slope of the associated equation (page 25, no number) is positive with respect to g. This would move the estimated value toward the observed (12% estimated and 8% observed), and (4) a general increase in the estimated values listed in Tables IV and V. This is based on observing that the slope of Eqs. (15) and (16) are negative with respect to g. This would move the estimated values toward the observed.

III. APPLICATION OF THE MODEL

A. DATA SOURCE

With the estimates of g and c being obtainable from only a knowledge of the total number of errors and the trial number of the last error, it was now a matter of obtaining some data to evaluate the model. The Institute has 41 programs in operation. These programs were being used in 20 different locations in the country. Due to the nature of the overall program, individual students receive only those programs which they require. It was therefore difficult to determine which program to sample. After some consideration the "What is" program was decided on. Calls were sent out to four sites for complete data on students having completed this program. There were 19 subjects from the 4 locations, the least from any location was 2 and the most was 7. There were 11 boys and 8 girls.

The "What is" program consists of 35 steps grouped into 9 series. With 19 subjects there was a possibility of obtaining (35x19) = 765 sets of data points, each representing a unique student/step combination. Throughout the remainder of this work "data point" will refer to a set of data from a unique student/step combination. Only 208 data points were available for the following reasons:

1. Due to the placement procedure the student may start at a point other than the beginning of the program. (There were 208 missing points due to this.)

2. If a student is put through a step and makes no errors, there is no way to tell whether he was in solution before he started or had an errorless response string. Therefore, these points yielded no usable information and had to be ignored. (There were 349 of these points.)

B. EXAMINATION OF DATA

The raw data formed a 35 by 19 matrix with 208 data points. Each data point consisted of the trial number of the last error (T) and the total number of errors (E). The data was then converted into a similar matrix containing $\hat{c}$ and $\hat{g}$ at each data point using Eqs. (11) and (12). The first step was to examine the various frequency distributions to see what information they would yield. Frequency distributions are presented for the number of trials (T), number of errors (E), $\hat{c}$ and $\hat{g}$. These distributions consisted of a total of 208 data points each and are shown in Figs. 1, 2, 3, and 4 respectively. Figure 1 shows the frequency distribution of the trial number of the last error, the interval widths are two trials. Figure 2 shows the frequency distribution of the number of errors, the interval widths are one error. Figures 3 and 4 show the frequency distribution of $\hat{c}$ and $\hat{g}$ respectively, the interval widths are .04. Figures 1, 2, and 3 exhibit a consistent decreasing nature, which indicates the lower values of T, E, and $\hat{c}$ are predominant in the sample and should generally be expected if the sample is representative of the population. Figure 4 has a much different shape, somewhat similar to a truncated normal type distribution. Figures 3 and 4 each have one point hat

breaks the otherwise consistent pattern. These points are in the interval .48 to .52 for $\hat{c}$, and .64 to .68 for $\hat{g}$. There appears to be a relatively simple explanation for these points: Data points having 1 error and 1 trial in presolution, that is an error on the first trial only, yields a $\hat{c}$ = .50, and a $\hat{g}$ = .667. If a student was in solution, i.e., knew how to answer correctly, but misunderstood the protocol the chances are good that he would have one error only. This protocol concept seems a valid explanation for the out of place peaks in Figs. 3 and 4.

Examination of the individual data (Appendixes A and B) showed no consistent trends. All of the factors T, E, $\hat{c}$ and $\hat{g}$ vary both over subjects and over steps. Accordingly the next approach was to calculate the means and standard deviations of $\hat{c}$ and $\hat{g}$ by subject and by step. The results are shown in Tables I and II, respectively. The frequency distributions of these values are plotted in Figs. 5 through 8. Figures 5 and 6 show $\hat{c}$ by subject and step, respectively. It is apparent that these frequency distributions overlap and that they fall under the overall frequency distribution of $\hat{c}$ shown in Fig. 3. Similarly Figs. 7 and 8 overlap and fall under the overall frequency distribution of $\hat{g}$ shown in Fig. 4.

It had been hoped that some definite consistency in one of the parameters would show up over subjects or steps. It appears, however, that the parameters g and c are random variables from the overall distributions shown in Figs. 3 and 4. Assuming that the sample is representative, it then follows that n this

TABLE I

MEAN AND STANDARD DEVIATION OF g AND c BY SUBJECT

	$\hat{g}$		$\hat{c}$	
Subject	Mean	Standard Deviation	Mean	Standard Deviation
1	.738	.167	.023	.019
2	.712	.283	.060	.040
3	.789	.119	.109	.149
4	.752	.143	.081	.065
5	.840	.038	.144	.077
6	.769	.174	.148	.154
7	.764	.085	.225	.240
8	--	--	--	--
9	.725	.245	.095	.105
10	.824	.099	.083	.139
11	.885	.135	.154	.214
12	.759	.119	.057	.063
13	.767	.144	.025	.023
14	.811	.136	.230	.248
15	.741	.146	.048	.044
16	.790	.128	.092	.132
17	.805	.105	.173	.197
18	.837	.101	.119	.191
19	.667	0	.500	0

TABLE II

MEAN AND STANDARD DEVIATION OF g AND c BY STEP

	$\hat{g}$		$\hat{c}$			$\hat{g}$		$\hat{c}$	
Step	Mean	Standard Deviation	Mean	Standard Deviation	Step	Mean	Standard Deviation	Mean	Standard Deviation
1	.554	.260	.183	.190	19	.819	.114	.063	.059
2	.667	.097	.315	.216	20	.736	.205	.063	.055
3	.635	.125	.268	.255	21	.797	.086	.067	.087
4	.723	.196	.189	.271	22	.760	.214	.070	.057
5	.792	.185	.085	.123	23	.794	.066	.020	.020
6	.876	.065	.085	.094	24	.719	.093	.136	.202
7	.758	.112	.062	.091	25	.764	.147	.092	.120
8	.812	.126	.182	.276	26	.775	.081	.041	.056
9	.852	.085	.051	.077	27	.821	.038	.054	.028
10	.884	.049	.065	.072	28	.835	.106	.091	.095
11	.809	.141	.036	.027	29	.882	.075	.043	.038
12	.774	.124	.205	.259	30	.842	.095	.072	.051
13	.764	.192	.053	.051	31	.777	.122	.136	.185
14	.746	.159	.131	.128	32	.747	.143	.027	.025
15	.742	.126	.021	.013	33	.851	.085	.068	.054
16	.852	.185	.054	.025	34	.855	.111	.093	.100
17	.767	.218	.040	.032	35	.834	.100	.037	.026
18	.772	.128	.045	.029					

program the values of g and c, and therefore L and E, are random variables from the general populations indicated by Figs. 1 through 4.

C. CRITERIA DETERMINATION

Based on this conclusion, general criteria for all students on this program can be determined. Two criteria have to be established: (1) At what point in the program can the student be advanced with a high probability of being in solution? (2) At what point in the program should a student be sent to a subroutine which has lower difficulty levels?

1. Advancement Criterion

The first criterion is based on the student's g level for that step and can be analyzed from the individual data given in Appendix B and Fig. 4. The individual data (Appendix B) showed that the largest $\hat{g}$ for this population is .974. Using this as an indication of the highest g that might be expected, the next step is to determine its approx. probability of occurrence [P(g=M)]. Figure 4 shows that there were 8 times when $\hat{g}$ was recorded between .96 and 1.0. This represents approximately .04 or 4 percent.

$$\therefore P[g=.98] \simeq .04$$

Table III gives an indication of what the probability would be of not being in solution. The values listed are for several different consecutive numbers of correct responses (X) as the criterion. M is the midpoint value of $\hat{g}$ used in Fig. 4, i.e., (.98, .94, .92.....) and is used to represent

TABLE III

ESTIMATED PROBABILITY OF NOT BEING IN SOLUTION [$P(\overline{S})$] ONCE THE CRITERION OF X CONSECUTIVE CORRECT RESPONSE HAS BEEN ACHIEVED

Criterion X	Estimate $P(\overline{S})$
8	.26
10	.21
12	.17
14	.14
16	.12
18	.10

the interval. The following basic, conditional probability statement was used for these calculations.

$$P[\text{Not Solution}]=P[\overline{S}]=\sum_{M} P[\text{Not Solution}\,|\,g=M,\text{Criterion}=X]\,P[g=M]$$

$$\therefore\ P[\overline{S}]=\sum_{M}(M)^{X}\cdot P[g=M]$$

It is obvious that the criterion must be established based on the desired certainty. For purposes of validation, let the criterion for completion be established at 16 consecutive correct responses. The actual raw data criterion levels of 10 and 20 consecutive correct responses were used in a rather arbitrary fashion. Of the possible 765 data points, there were 765 - 208 = 557 in which the subject was tested. Of these 557 data points, there were 306 which used a criterion level of 20. Of these 306 data points, there

were only 24 in which a criterion level of 16 would not have been satisfactory. That is, there were 24 data points where 16, 17, 18, or 19 consecutive correct responses were recorded followed by at least one more errors. This value

$$\frac{24}{306} \simeq .08 = 8\%$$

means that 8 percent of the time the criterion of 16 was inadequate. The prediction was 12 percent, and therefore the method of setting this first criterion level seems appropriately conservative.

2. Branching Criterion

To establish the second criterion we turn to Figs. 1 and 2. Figure 2 tells us that 90% of the data points had less than 25 errors. Figure 1 tells us that 87% of the data points were completed in less than 90 trials. What is more significant about Fig. 1 is that the data points which had more than 90 trials required a total of 5065 trials, whereas those requiring less than 90 trials required only a total of 3988 trials. This means that 13% of the data points required approximately 56% of the effort. Similarly, the upper 10% of the data points showed 1586 errors out of a total of 2575 error. This means that approximately 61% of all the errors occurred in 10% of the data point. These two observations indicate that there should be some upper value of E and/or T which would signify the need for branching. That is, some point that, if reached, would indicate that the subject is having an exceptional amount of difficulty and should be branched. To develop this further

requires returning to Eqs. (5) and (8). The following shows that these equations are true probability mass functions.

$$\lim_{k\to\infty} P[E\leq k] = \sum_{i=0}^{\infty} \rho(1-\rho)^{i}$$

$$= \frac{\rho}{1-(1-\rho)} = 1$$

$$\lim_{t\to\infty} P[T\leq t] = \rho + \sum_{i=1}^{\infty} \rho(1-g)(1-c)^{i}$$

$$= \rho+\rho(1-g)(1-c)\sum_{i=1}^{\infty}(1-c)^{i-1}$$

$$= \rho+\rho(1-g)(1-c)\frac{1}{c}$$

$$= \rho\left(1+\frac{(1-g)(1-c)}{c}\right)$$

$$= \rho\left(\frac{c+1-c-g(1-c)}{c}\right)$$

$$= \rho\left(\frac{1-g(1-c)}{c}\right)$$

$$= \rho\left(\frac{1}{\rho}\right) = 1$$

Based on this development it follows that

$$P[E>k] = 1 - \sum_{i=0}^{k} \rho(1-\rho)^{i}$$

$$= 1 - \frac{\rho[1-(1-\rho)^{k+1}]}{1-(1-\rho)}$$

$$= 1 - 1 + (1-\rho)^{k+1}$$

$$\therefore P[E>k] = \left[\frac{(1-g)(1-c)}{1-g(1-c)}\right]^{k+1} \quad (13)$$

and

$$P[T>t] = 1-\rho-\sum_{i=1}^{t}\rho(1-g)(1-c)^{i}$$

$$= (1-\rho)-\rho(1-g)(1-c)\sum_{i=1}^{t}(1-c)^{i-1}$$

$$= (1-\rho)-\rho(1-g)(1-c)\frac{[1-(1-c)^{t}]}{c}$$

$$= \frac{(1-g)(1-c)}{1-g(1-c)} - \frac{(1-g)(1-c)}{1-g(1-c)}[1-(1-c)^{t}]$$

$$\therefore P[T>t] = \frac{(1-g)(1-c)}{1-g(1-c)}[1-c]^{t} \tag{14}$$

With Eqs. (13) and (14) in hand, it simply becomes a matter of calculations. The calculations are similar to those done for the first criterion. To get the unconditional probability that E or T is greater than some number requires the following basic type equations, involving two independent random variables (g and c).

$$P[E>k] = \sum_{g}\sum_{c} P[E>k|g=x,c=y]\ P[g=x]\ P[c=y]$$

$$= \sum_{g}\sum_{c} \left[\frac{(1-x)(1-y)}{1-x(1-y)}\right]^{k+1} P[g=x]\ P[c=y] \tag{15}$$

and

$$P[T>t] = \sum_{g}\sum_{c} P[L>t|g=x,c=y]\ P[g=x]\ P[c=y]$$

$$= \sum_{g}\sum_{c} \frac{(1-x)(1-y)}{1-x(1-y)}[1-y]^{t}\ P[g=x]\ P[c=y] \tag{16}$$

Tables IV and V represent the results for E and T respectively. Also listed in Tables IV and V are the values obtained om the

lower abscissa's of Figs. 2 and 1, respectively. The lower abscissa's represent the cumulative percentages, that is, the percentage of data points with values less than or equal to the value indicated on the first abscissa. For the sake of comparison, one minus the values in Figs. 2 and 1 are used in Tables IV and V. This value then represents the percentage of data points in which E and T were greater than the associated criterion number.

TABLE IV

ESTIMATED PROBABILITY THAT THE NUMBER OF ERRORS WILL EXCEED A GIVEN CRITERION k [P(E>k)]

Criterion k	Estimate P(E>k)	Observed Figure 2
5	.29	.36
10	.17	.27
15	.12	.20
20	.08	.14
25	.05	.10

(Also associated with this table are the observed percentages of data points which had more than k errors [from Fig. 2]).

It should be noted that the observed values in Tables IV and V are consistently higher than the expected values. Some thought indicated that the assumption of independence, required for the development of Eqs. (15) and (16) may n t have held up. A Pearson "r" test was run with g=X and [illegible] Y.

TABLE V

ESTIMATED PROBABILITY THAT THE NUMBER OF TRIALS WILL EXCEED A GIVEN CRITERION t [P(T>t)]

Criterion t	Estimate P(T>t)	Observed Figure 1
60	.13	.20
70	.10	.18
80	.08	.16
90	.07	.13
100	.06	.12

(Also associated with this table are the observed percentages of data points which had more than t trials [from Fig. 1]).

The result was r=-.2774 with n=208. This value is significant at greater than the P=.01 level. With a knowledge of this fact an examination of Eqs. (15) and (16) indicates why the estimated values were lower than the observed values. It is felt however that the correlation between $\hat{g}$ and $\hat{c}$ is an unavoidable artifact of the estimators. In explanation of this last statement, observe that Eqs. (11) and (12) have an upper and lower bound based on the value of T. These bounds are most restrictive at the lower values of T where $\hat{c}$ is larger than $\hat{g}$ has its most restrictive upper bound. Since 36% of the population had a T value of ten or less (Fig. 1), it is relatively easy to accept a higher $\hat{c}$ and lower $\hat{g}$ (negative r) bias in the estimators. The branching criterion must therefore be established with an awareness of this artifact.

One case of branching was observed in the raw data. At that time the branching criterion required three consecutive sessions with a percent of accuracy less than 80%. It took 184 trials before branching was performed. Had a criterion of 25 errors been used, the subject would have branched after the 57th trial, which would obviously have been an advantageous move. Since there is no apparent way to measure the benefit of branching over just forcing the subject through the step, a comparison similar to that done for the first criterion cannot be made. It is felt however that the confidence on using the number of errors or the number of trials as branching criterion will lie somewhere between the estimated and the observed values (Tables IV and IV).

IV. CONCLUSION

In general, any program that has the form of those used in this study could be analyzed in a similar manner. Once the program is ready for testing, a representative sample should be run through with an arbitrarily high criterion, say 20 consecutive correct responses with no branching. (Branching changes the basic probabilities (g and c) and confounds the data.) Analysis of the resulting data would then indicate the criteria to be used when the program is given general use. For this particular program the following criteria seem adequate:

Pass Criterion = 16 consecutive correct responses (88-92% confidence)

Branch Criterion = 25 total errors (90-95% confidence)

or Branch Criterion = 100 total trials (88-94% confidence)

It is recommended that anyone using this procedure attempt to increase the number of data points at least twofold. This increased number of data points would smooth out the frequency distribution and possibly point more exactly to how c and g are distributed. Having a smoother frequency distribution will also fill in blanks in the joint probability matrix used in calculating $P(E>k)$ and $P(T>t)$. With more cells filled in, it may be possible to use the actual joint density $P(g=X, c=Y)$ rather than the product of the marginal densities $P(g=X) \cdot P(c=Y)$. This would bring the estimated and observed values in Tables IV and V closer together because it would remove the requirement for an assumption of independence which appears to

violated due to the model. Knowledge of this procedure before the data is taken will also allow for better control of the data gathering.

APPENDIX A

Individual Data Matrix (T and E)
Rows Represent Subjects, Columns Represent Steps
Top Entry Equals T, Bottom Entry Equals E

Subject	A_1	A_2	A_3	A_4	A_5	A_6	A_7	B_1	B_2	B_3	B_4	B_5
1	12	71	47	0	0	0	0	0	0	0	0	0
	4	42	18	0	0	0	0	0	0	0	0	0
2	7	0	0	0	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	0	0	0	0	0
3	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-
5	0	0	9	0	0	0	0	0	0	0	0	0
	0	0	3	0	0	0	0	0	0	0	0	0
6	0	2	1	0	13	0	0	0	0	0	0	0
	0	2	1	0	1	0	0	0	0	0	0	0
7	-	-	-	-	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-	-	-	-	-
9	4	0	0	0	2	0	0	0	0	0	0	0
	4	0	0	0	2	0	0	0	0	0	0	0
10	0	1	1	8	8	11	76	108	219	44	99	62
	0	1	1	1	1	1	13	15	41	5	25	17

APPENDIX A

Individual Data Matrix (T and E)
Rows Represent Subjects, Columns Represent Steps
Top Entry Equals T, Bottom Entry Equals E

Subject	C_1	C_2	C_3	C_4	C_5	D_1	D_2	D_3	D_4	D_5	E_1	E_2
1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	39	0	14	10	0	0	0	24	0	0
	0	0	12	0	2	2	0	0	0	5	0	0
3	-	-	-	-	-	37	0	0	8	0	0	1
						14	0	0	3	0	0	1
4	5	3	22	0	0	9	0	0	15	25	0	0
	3	1	8	0	0	5	0	0	4	13	0	0
5	0	0	0	0	0	0	3	39	2	4	0	0
	0	0	0	0	0	0	1	5	1	1	0	0
6	39	2	0	12	0	0	0	0	22	0	0	0
	3	2	0	7	0	0	0	0	3	0	0	0
7	-	-	-	-	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-	-	-	-	-
9	41	0	0	0	0	14	3	41	0	6	0	0
	7	0	0	0	0	1	1	18	0	5	0	0
10	6	0	43	30	0	26	85	15	221	0	16	2
	2	0	5	4	0	6	14	2	45	0	3	1

APPENDIX A

Individual Data Matrix (T and E)

Rows Represent Subjects, Columns Represent Steps

Top Entry Equals T, Bottom Entry Equals E

Subject	E_3	E_4	E_5	F_1	F_2	F_3	G_1	G_2	G_3	H_1	I_1
1	0 0	0 0	0 0	0 0	0 0	0 0	109 50	0 0	0 0	54 4	-
2	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
3	0 0	0 0	9 2	11 2	0 0	9 4	0 0	69 20	8 1	0 0	8 1
4	35 5	0 0	27 5	0 0	7 1	8 1	0 0	10 6	23 6	2 1	23 7
5	0 0	0 0	0 0	0 0	17 3	3 1	3 1	0 0	3 1	2 1	0 0
6	10 3	0 0	0 0	2 1	4 1	0 0	10 4	35 6	0 0	12 1	-
7	-	-	-	-	0 0	0 0	1 1	12 3	6 2	-	-
8	-	-	-	-	-	-	0 0	0 0	0 0	-	-
9	0 0	0 0	0 0	0 0	0 0	0 0	0 0	55 5	47 5	16 2	0 0
10	2 2	57 13	0 0	5 1	43 8	37 5	0 0	310 79	129 30	16 3	-

APPENDIX A

Individual Data Matrix (T and E)

Rows Represent Subjects, Columns Represent Steps

Top Entry Equals T, Bottom Entry Equals E

Subject	A_1	A_2	A_3	A_4	A_5	A_6	A_7	B_1	B_2	B_3	B_4	B_5
11	0	1	1	0	0	0	0	0	0	0	15	0
	0	1	1	0	0	0	0	0	0	0	1	0
12	0	3	0	0	91	48	178	0	3	3	0	0
	0	2	0	0	34	9	67	0	1	1	0	0
13	0	32	138	277	12	18	0	21	59	8	13	5
	0	8	85	123	2	2	0	3	15	1	5	1
14	0	1	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0
15	-	-	-	-	-	-	-	-	-	-	-	-
16	0	14	0	1	96	9	2	0	38	106	0	0
	0	3	0	1	19	2	1	0	2	19	0	0
17	1	1	0	0	0	2	0	0	0	0	0	0
	1	1	0	0	0	1	0	0	0	0	0	0
18	44	1	0	0	0	0	0	0	0	0	0	0
	10	1	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	1	0	0	0	1
	0	0	0	0	0	0	0	1	0	0	0	1

APPENDIX A

Individual Data Matrix (T and E)

Rows Represent Subjects, Columns Represent Steps

Top Entry Equals T, Bottom Entry Equals E

Subject	C_1	C_2	C_3	C_4	C_5	D_1	D_2	D_3	D_4	D_5	E_1	E_2
11	10	0	0	0	0	0	0	0	0	19	0	0
	1	0	0	0	0	0	0	0	0	1	0	0
12	20	36	269	6	212	0	0	0	334	0	262	55
	10	10	99	1	104	0	0	0	119	0	57	17
13	170	0	28	16	0	233	53	37	0	200	133	54
	93	0	4	1	0	78	24	9	0	39	40	14
14	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
15	-	-	-	-	-	27	21	6	28	0	43	78
						6	3	5	4	0	9	16
16	0	7	0	0	0	11	0	0	0	4	0	52
	0	3	0	0	0	3	0	0	0	1	0	24
17	0	0	0	9	10	0	15	0	0	0	0	0
	0	0	0	1	2	0	4	0	0	0	0	0
18	0	0	0	0	0	0	23	4	0	0	0	0
	0	0	0	0	0	0	3	1	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX A

Individual Data Matrix (T and E)

Rows Represent Subjects, Columns Represent Steps

Top Entry Equals T, Bottom Entry Equals E

Subject	E_3	E_4	E_5	F_1	F_2	F_3	G_1	G_2	G_3	H_1	I_1
11	0	0	0	0	17	0	21	0	8	0	0
	0	0	0	0	1	0	2	0	1	0	0
12	16	101	12	0	142	7	9	13	4	0	0
	4	15	3	0	32	1	3	3	1	0	0
13	12	80	49	47	178	104	23	64	334	0	88
	2	18	11	17	44	19	2	24	112	0	17
14	0	0	0	0	9	0	1	86	6	0	0
	0	0	0	0	1	0	1	14	1	0	0
15	70	6	9	37	0	9	13	184	0	82	0
	22	3	3	4	0	3	4	86	0	32	0
16	0	0	0	0	22	0	0	50	0	0	0
	0	0	0	0	2	0	0	23	0	0	0
17	0	0	0	0	0	0	36	41	0	6	0
	0	0	0	0	0	0	6	13	0	1	0
18	0	0	0	0	0	0	75	36	0	0	0
	0	0	0	0	0	0	15	3	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0

APPENDIX B

Individual Data Matrix ($\hat{g}$ and $\hat{c}$)
Rows Represent Subjects, Columns Represent Steps
Top Entry Equals $\hat{g}$, Bottom Entry Equals $\hat{c}$
(S = Tested with No Errors, S' = Not Tested)

Subject	A_1	A_2	A_3	A_4	A_5	A_6	A_7	B_1	B_2	B_3	B_4	B_5
1	.750 .067	.570 .014	.644 .020	S	S	S	S	S	S	S	S	S
2	.222 .125	S	S	S	S	S	S	S	S	S	S	S
3	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'
4	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'
5	S	S	.770 .083	S	S	S	S	S	S	S	S	S
6	S	.500 .333	.667 .500	S	.963 .038	S	S	S	S	S	S	S
7	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'
8	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'
9	.333 .200	S	S	S	.500 .333	S	S	S	S	S	S	S
10	S	.667 .500	.667 .500	.941 .063	.941 .063	.956 .045	.843 .012	.871 .009	.818 .004	.907 .019	.760 .010	.745 .015

APPENDIX B

Individual Data Matrix ($\hat{g}$ and $\hat{c}$)
Rows Represent Subjects, Columns Represent Steps
Top Entry Equals $\hat{g}$, Bottom Entry Equals $\hat{c}$
(S = Tested with No Errors, S' = Not Tested)

Subject	C_1	C_2	C_3	C_4	C_5	D_1	D_2	D_3	D_4	D_5	E_1	E_2
1	.927	.911	.610	S	S	S	S	S	S	S	S	S
	.025	.023	.008									
2	S	S	.723	S	.909	.875	S	S	S	.832	S	S
			.024		.048	.067				.035		
3	S'	S'	S'	S'	S'	.656	S	S	.743	S	S	.667
						.025			.094			.500
4	.609	.857	.689	S	S	.576	S	S	.797	.534	S	S
	.150	.167	.040			.093			.053	.037		
5	S	S	S	S	S	S	.857	.845	.800	.889	S	S
							.167	.021	.250	.125		
6	.943	.500	S	.524	S	S	S	S	.901	S	S	S
	.019	.333		.073					.034			
7	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'
8	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'
9	.854	S	S	S	S	.965	.857	.593	S	.390	S	S
	.021					.036	.125	.023		.139		
10	.800	S	.905	.896	S	.809	.848	.915	.802	S	.866	.800
	.111		.019	.027		.033	.011	.044	.004		.047	.250

APPENDIX B

Individual Data Matrix ($\hat{g}$ and $\hat{c}$)
Rows Represent Subjects, Columns Represent Steps
Top Entry Equals $\hat{g}$, Bottom Entry Equals $\hat{c}$
(S = Tested with No Errors, S' = Not Tested)

Subject	E_3	E_4	E_5	F_1	F_2	F_3	G_1	G_2	G_3	H_1	I_1
1	S	S	S	S	S	S	.554 .009	S	S	.941 .015	S'
2	S	S	S	S	S	S	S	S	S	S	S
3	S	S	.862 .074	.886 .061	S	.673 .089	S	.728 .014	.941 .063	S	.941 .063
4	.884 .024	S	.850 .031	S	.933 .071	.941 .063	S	.526 .086	.784 .037	.800 .250	.743 .038
5	S	S	S	S	.873 .044	.857 .167	.857 .167	S	.857 .167	.800 .250	S
6	.791 .075	S	S	.800 .250	.889 .125	S	.704 .080	.857 .024	S	.960 .042	S'
7	S'	S'	S'	S'	S	S	.667 .500	.834 .063	.800 .111	S'	S'
8	S'	S'	S'	S'	S'	S'	S	S	S	S'	S'
9	S	S	S	S	S	S	S	.925 .015	.913 .018	.920 .042	S
10	.500 .333	.792 .016	S	.909 .100	.838 .021	.890 .023	S	.749 .003	.777 .008	.866 .047	S'

APPENDIX B

Individual Data Matrix ($\hat{g}$ and $\hat{c}$)

Rows Represent Subjects, Columns Represent Steps

Top Entry Equals $\hat{g}$, Bottom Entry Equals $\hat{c}$

(S = Tested with No Errors, S' = Not Tested)

Subject	A_1	A_2	A_3	A_4	A_5	A_6	A_7	B_1	B_2	B_3	B_4	B_5
11	S	.667 .500	.667 .500	S	S	S	S	S	S	S	.968 .033	S
12	S	.636 .222	S	S	.641 .011	.834 .019	.631 .006	S	.857 .167	.857 .167	S	S
13	S	.784 .028	.396 .007	.561 .004	.895 .056	.929 .037	S	.897 .036	.765 .016	.941 .063	.699 .064	.909 .100
14	S	.667 .500	S	S	S	S	S	S	S	S	S	S
15	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'	S'
16	S	.847 .054	S	.667 .500	.814 .010	.862 .074	.800 .167	S	.966 .018	.831 .009	S	S
17	.667 .500	.667 .500	S	S	S	.800 .250	S	S	S	S	S	S
18	.798 .021	.667 .500	S	S	S	S	S	S	S	S	S	S
19	S	S	S	S	S	S	S	.667 .500	S	S	S	.667 .500

APPENDIX B

Individual Data Matrix ($\hat{g}$ and $\hat{c}$)

Rows Represent Subjects, Columns Represent Steps

Top Entry Equals $\hat{g}$, Bottom Entry Equals $\hat{c}$

(S = Tested with No Errors, S' = Not Tested)

Subject	C_1	C_2	C_3	C_4	C_5	D_1	D_2	D_3	D_4	D_5	E_1	E_2
11	.952 .050	S	S	S	S	S	S	S	S	.974 .026	S	S
12	.565 .045	.754 .025	.637 .004	.923 .083	.516 .005	S	S	S	.648 .003	S	.787 .004	.713 .017
13	.462 .006	S	.889 .029	.970 .031	S	.671 .004	.573 .018	.786 .024	S	.811 .005	.709 .007	.762 .017
14	S	S	S	S	S	S	S	S	S	S	S	S
15	S'	S'	S'	S'	S'	.815 .032	.897 .036	.390 .139	.889 .029	S	.815 .021	.809 .012
16	S	.710 .107	S	S	S	.809 .068	S	S	S	.889 .125	S	.565 .018
17	S	S	S	.947 .056	.875 .067	S	.797 .053	S	S	S	S	S
18	S	S	S	S	S	S	.905 .033	.889 .125	S	S	S	S
19	S	S	S	S	S	S	S	S	S	S	S	S

APPENDIX B

Individual Data Matrix ($\hat{g}$ and $\hat{c}$)
Rows Represent Subjects, Columns Represent Steps
Top Entry Equals $\hat{g}$, Bottom Entry Equals $\hat{c}$
(S = Tested with No Errors, S' = Not Tested)

Subject	E_3	E_4	E_5	F_1	F_2	F_3	G_1	G_2	G_3	H_1	I_1
11	S	S	S	S	.971	S	.938	S	.941	S	S
					.029		.032		.063		
12	.810	.862	.824	S	.783	.933	.770	.836	.889	S	S
	.050	.009	.063		.007	.071	.083	.058	.125		
13	.895	.789	.798	.665	.760	.828	.944	.645	.689	S	.819
	.056	.012	.019	.020	.005	.009	.029	.015	.003		.011
14	S	S	S	S	.947	S	.667	.850	.923	S	S
					.056		.500	.011	.083		
15	.703	.667	.770	.915	S	.770	.768	.540	S	.626	S
	.014	.125	.083	.022		.083	.062	.005		.012	
16	S	S	S	S	.941	S	S	.567	S	S	S
					.030			.019			
17	S	S	S	S	S	S	.861	.712	S	.923	S'
							.024	.023		.083	
18	S	S	S	S	S	S	.815	.949	S	S	S
							.013	.021			
19	S	S	S	S	S	S	S	S	S	S	S

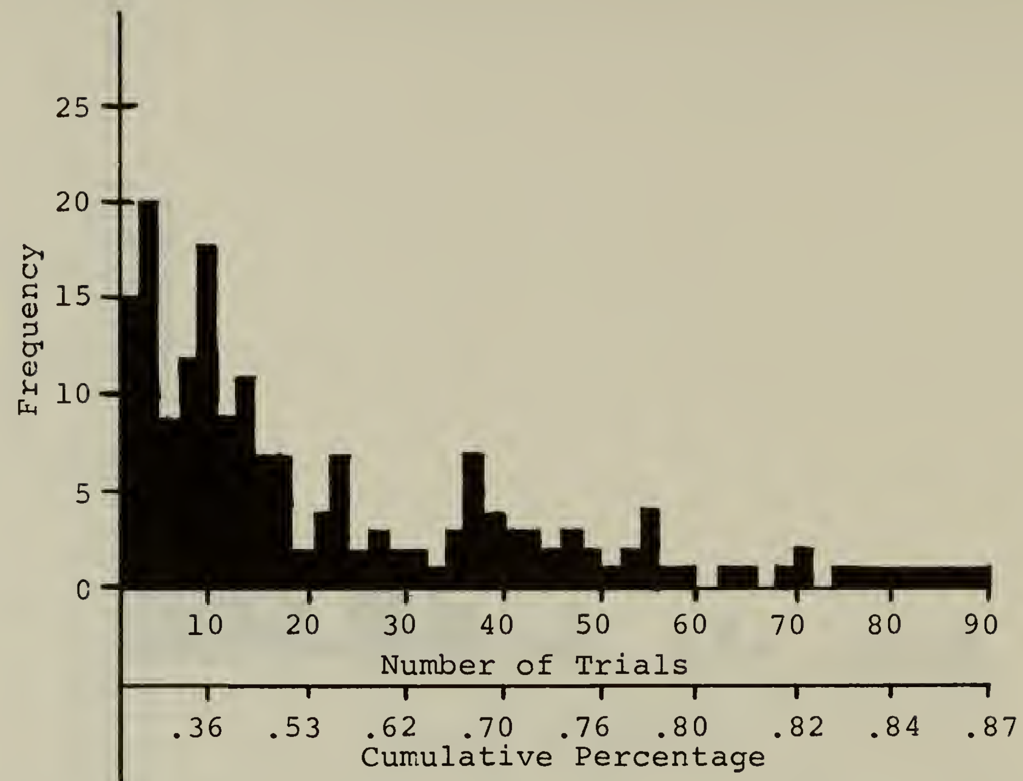

Figure 1. Frequency Distribution of The Total Number of Trials Prior to and Including the Last Error (T). (The values of T are listed on the first abscissa and the Cumulative Percentage are listed on the second abscissa. There where 28 occurrences of T greater than 90, ranging from 91 to 334. All occurred once except for T=334 which occurred twice.)

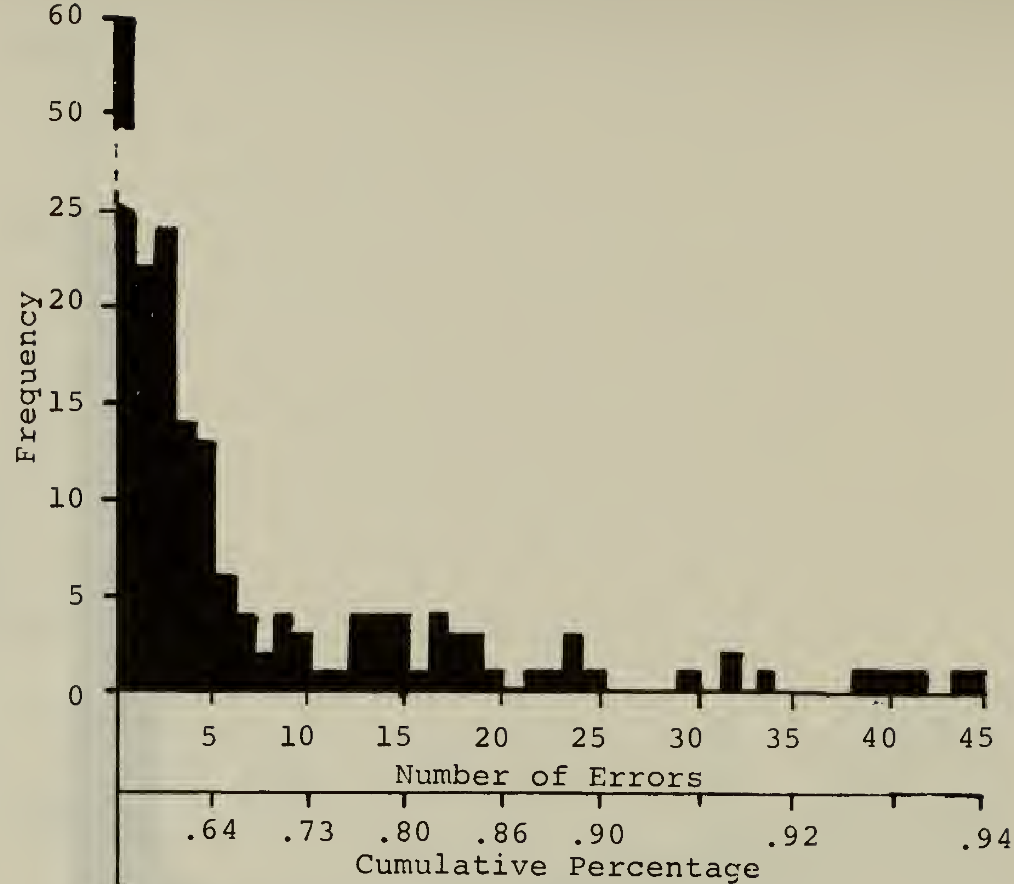

Figure 2. Frequency Distribution of The Total Number of Errors Prior To and Including The Last Error (E). (The values of E are listed on the first abscissa and the Cumulative Percentages are listed on the second abscissa. There were 14 occurrences of E greater than 45 ranging from 50 to 123 with no multiple occurrences.)

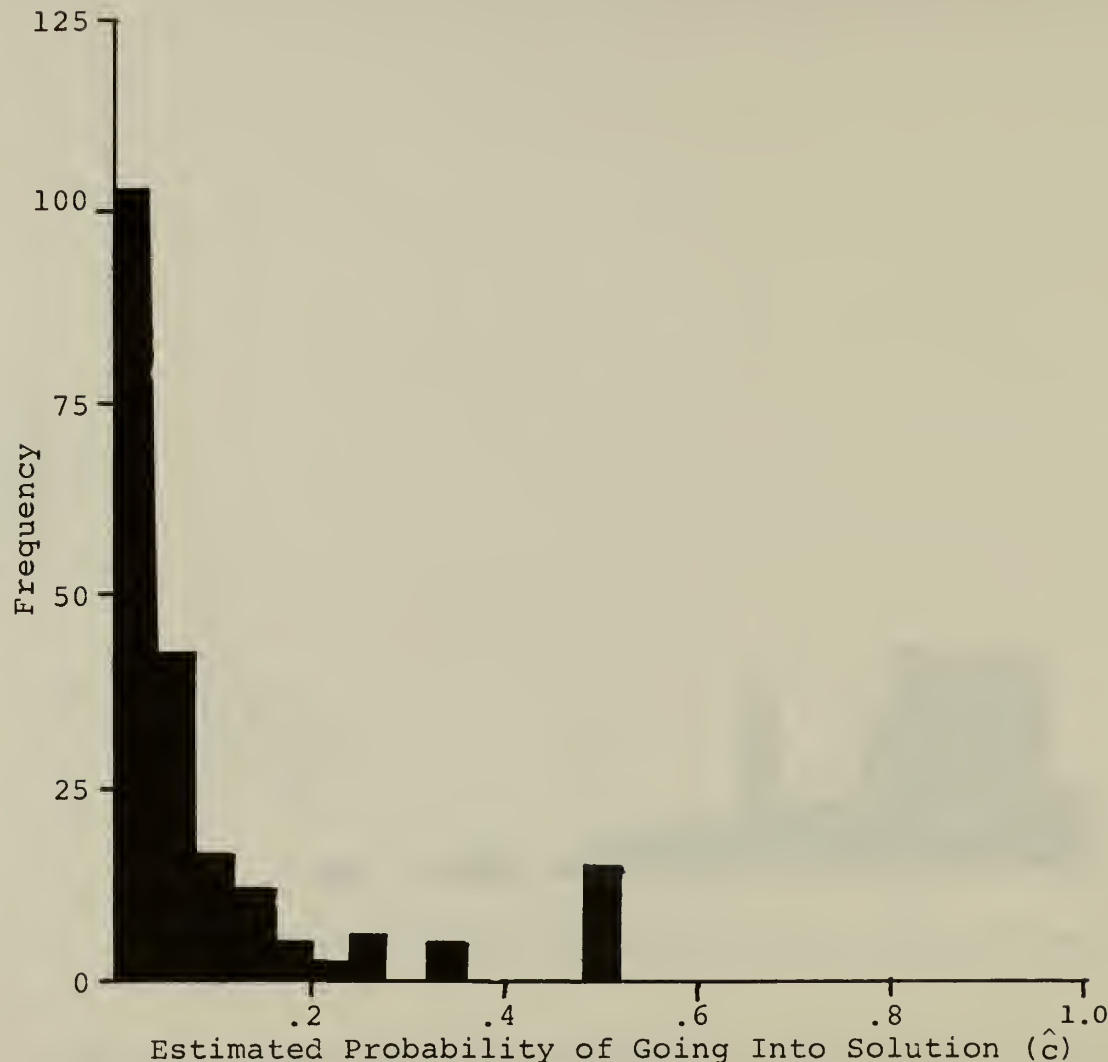

Figure 3. Frequency Distribution of the Estimated Probability of Going Into Solution ($\hat{c}$).

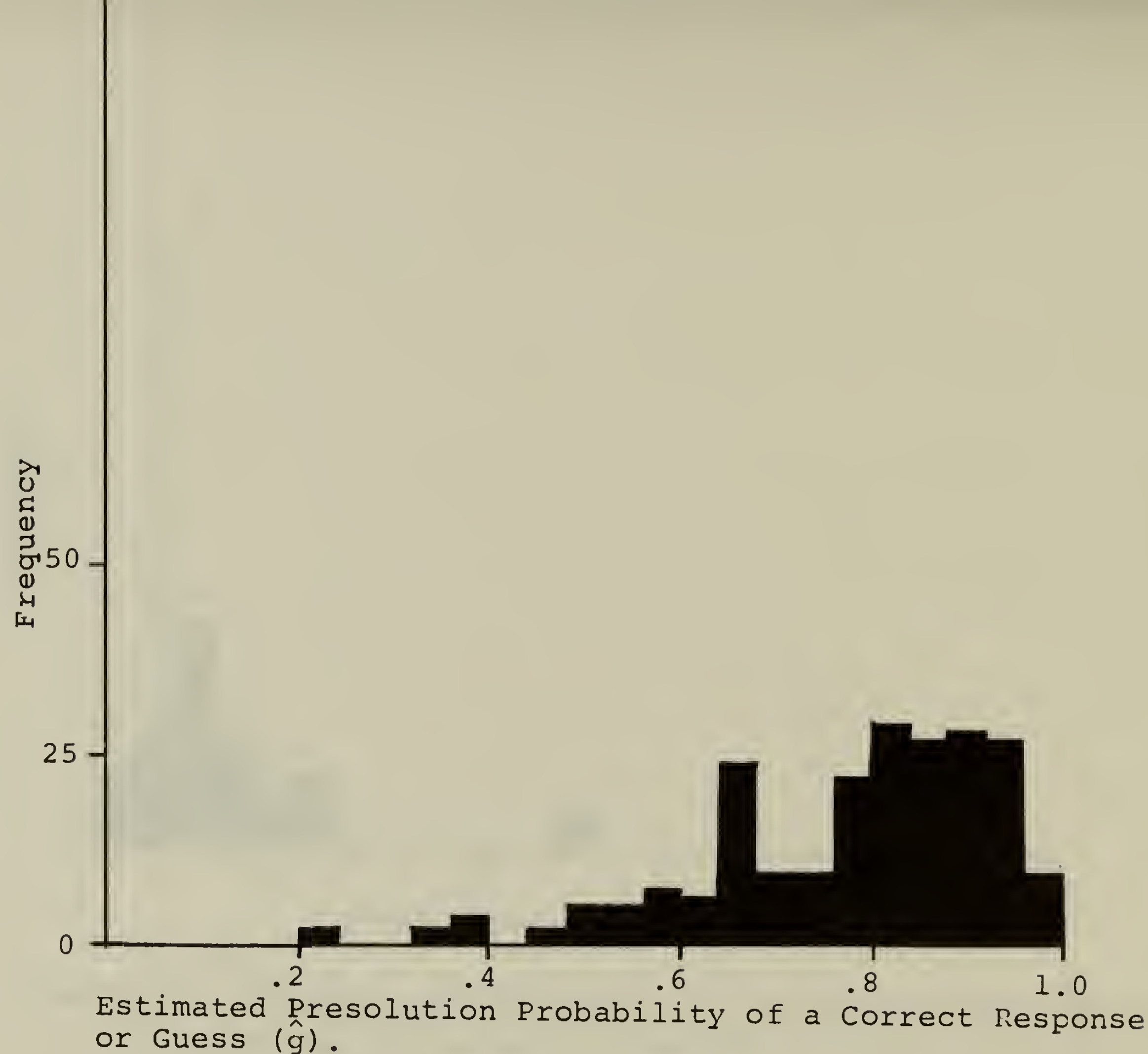

Figure 4. Frequency Distribution of The Estimated Presolution Probability of a Correct Response or Guess ($\hat{g}$).

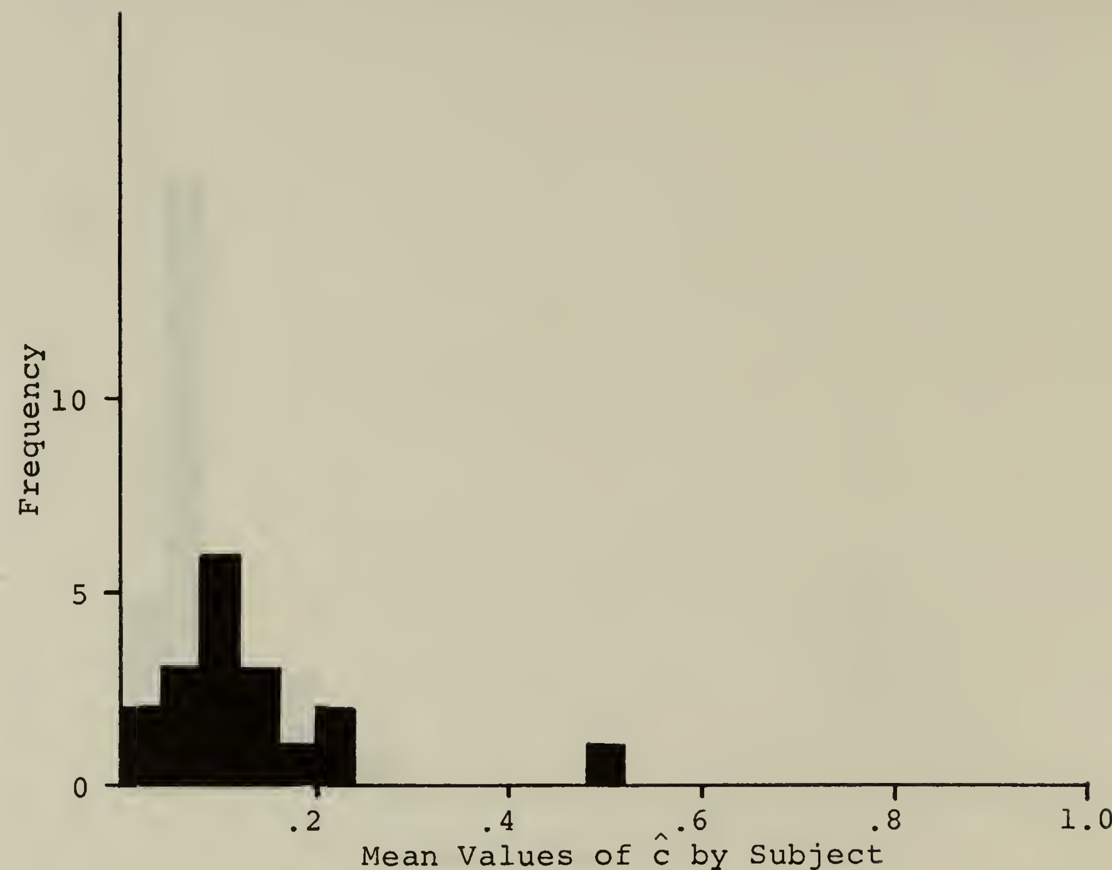

Figure 5. Frequency Distribution of The Mean Values (BY SUBJECT) of The Estimated Probability of Going Into Solution ($\hat{c}$).

Figure 6. Frequency Distribution of The Mean Values (BY STEP) of The Estimated Probability of Going Into Solution ($\hat{c}$).

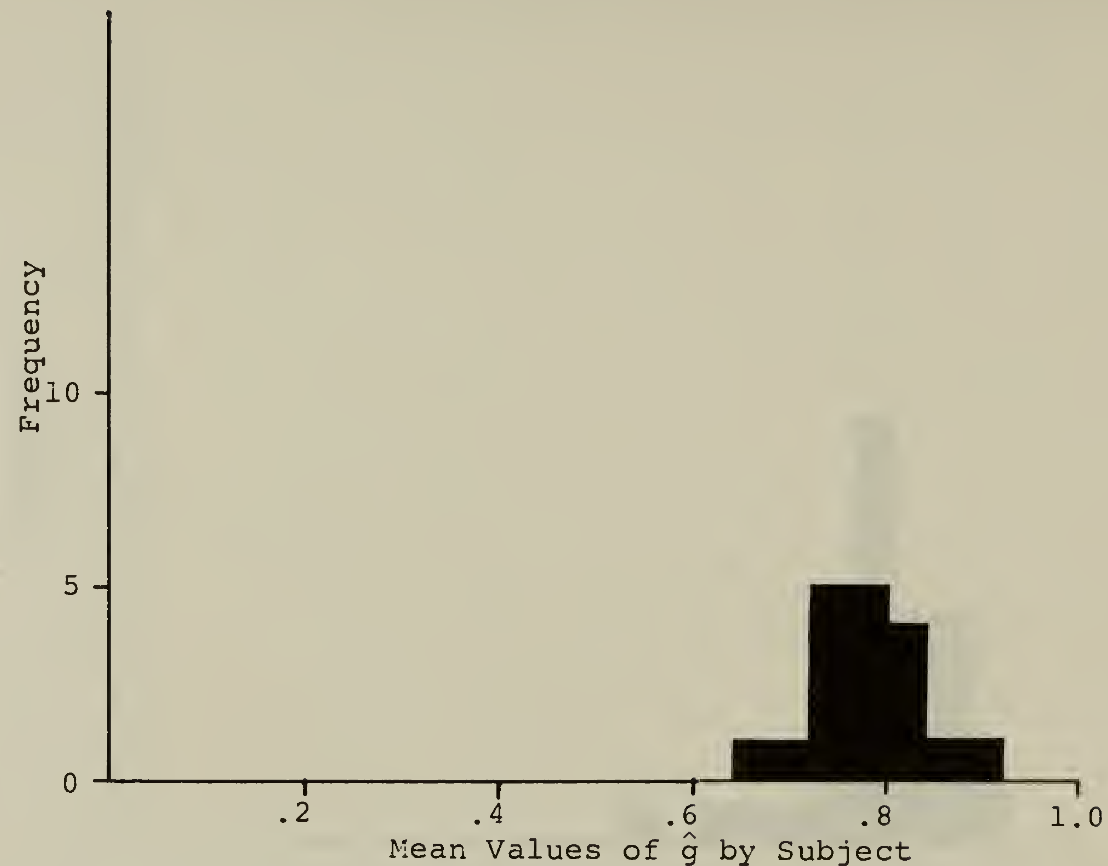

Figure 7. Frequency Distribution of The Mean Values (BY SUBJECT) of The Estimated Presolution Probability of a Correct Response or Guess ($\hat{g}$).

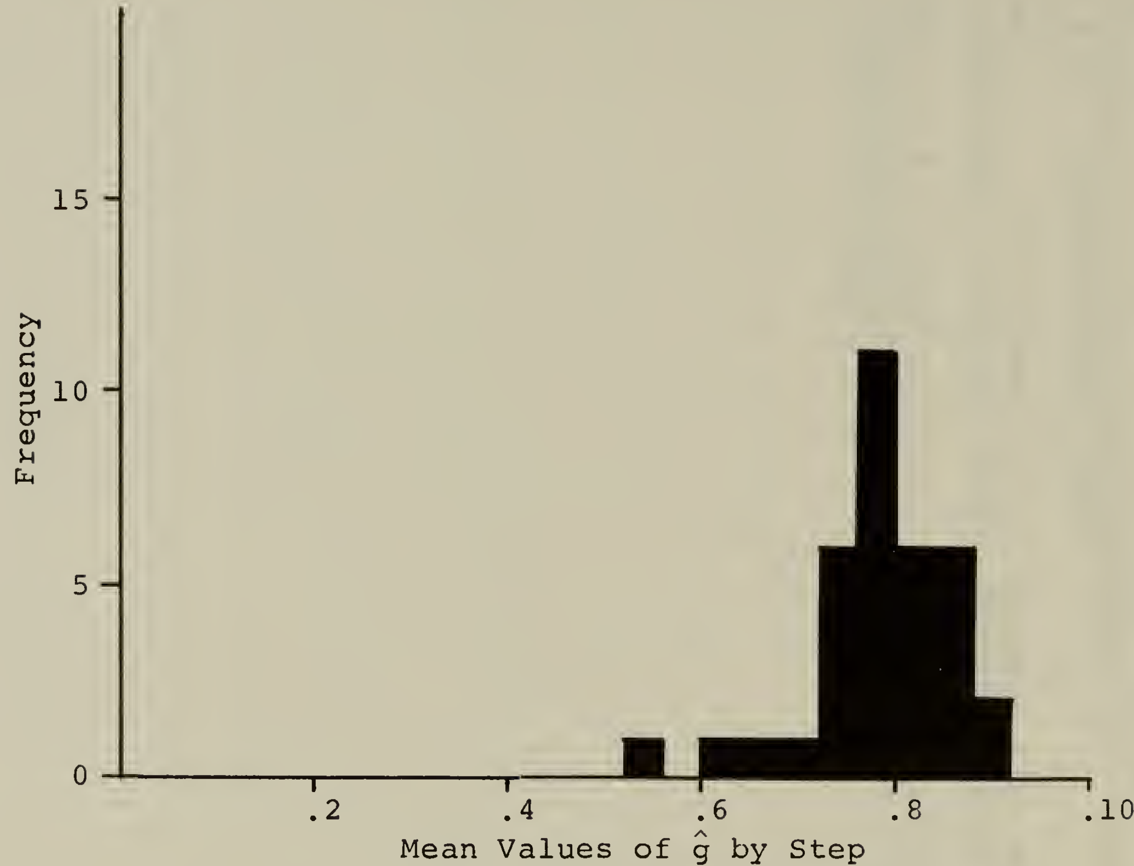

Figure 8. Frequency Distribution of The Mean Values (BY STEP) of The Estimated Presolution Probability of a Correct Response or Guess ($\hat{g}$).

LIST OF REFERENCES

1. Atkinson, R. C.; Bower, G. H.; and Crothers, E. J. An Introduction To Mathematical Learning Theory. New York: Wiley, 1965.

2. Coombs, C. H.; Dawes, R. M.; and Tversky, A. Mathematical Psychology An Elementary Introduction. Englewood Cliffs, N.J.: Prentice-Hall, 1970.

3. Restle, F., and Greene, J. G. Introduction To Mathematical Psychology. Reading, Mass: Addison-Wesley, 1970.

4. Gray, B. B., and Ryan, B. P. Language Training Through Programmed Conditioning. Champaign, Ill.: Research Press (In Press), 1973.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Assoc. Professor James K. Arima, Code 55Aa Department of Operations Research Naval Postgraduate School Monterey, California 93940	5
4.	CAPT G. G. Madson Headquarters U.S. Marine Corps Code AAM Washington, D.C. 20380	1
5.	Behavior Science Institute 969 Pacific Street Monterey, California 93940	1
6.	Assoc. Professor T. D. Burnett, Code 55Za Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
7.	Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
8.	Commanding General U.S.A. Combat Developments Command Experimentation Command Fort Ord, California 93941	1
9.	U.S.A. Behavioral and Systems Research Laboratory Room 239, The Commonwealth Building 1320 Wilson Boulevard Arlington, Virginia 22209	1

10. Commanding General 1
U.S.A. Human Engineering Laboratories
Aberdeen Proving Ground, Maryland 21005

11. Commanding General 1
U.S.A. Medical Research Laboratory
Fort Knox, Kentucky 40121

12. U.S. Air Force Human Resources Laboratory 1
Brooks Air Force Base, Texas 78235

13. Director 1
Personnel Research Laboratory
Lackland Air Force Base, Texas 78236

14. Engineering Psychology Division 1
Naval Training Device Center
Orlando, Florida 32013

15. Naval Aerospace Medical Institute 1
Pensacola, Florida 32512

16. Navy Medical Neuropsychiatric Research Unit 1
San Diego, California 92133

17. Naval Medical Research Laboratory 1
Submarine Base
New London, Connecticut 06342

18. Psychology Research Unit 1
Australian Military Forces
Albert Parks Barracks
Melbourne, Australia

19. Director 1
Human Resources Research Organization
300 North Washington Street
Alexandria, Virginia 22314

20. Training Research Distributor 1
U.S. Air Force Human Resources Laboratory
Wright-Patterson AFB, Ohio 45433

21. U.S. Naval Personnel Research
and Development Laboratory 1
Washington, D.C. 20390

22. U.S. Naval Personnel and Training Research 1
Laboratory
San Diego, California 92152

23. Commander (Code 4011) 1
Naval Weapons Center
China Lake, California 93555

24. Human Performance Branch 1
NASA Ames Research Center
Moffett Field, California 94035

25. Director 1
Human Resources Research Organization
P.O. Box 5787
Presidio of Monterey
Monterey, California 93940

26. Human Engineering Division 1
Aerospace Medical Laboratories
Wright-Patterson Air Force Base
Attn: MRHF
Dayton, Ohio 45433

27. Department of the Air Force 1
Department of Psychology and Leadership
U.S.A.F. Academy, Colorado 80840

28. Director, Office of Military Psychology 1
and Leadership
West Point, New York 10996

29. Commanding Officer 1
Naval Medical Field Research Laboratory
Camp Lejeune, North Carolina 28542

30. U.S. Naval Missile Test Center 1
Attn: Human Factors Branch
Point Mugu, California 93041

31. Psychological Sciences Division (Code 450) 1
Office of Naval Research
Department of the Navy
Arlington, Virginia 22217

32. U.S. Army Enlisted Evaluation Center 1
Fort Benjamin Harrison
Indianapolis, Indiana 46249

33. Head, Human Engineering Laboratory 1
Naval Electronics Laboratory Center
San Diego, California 92152

34. Head, Aerospace Operations Psychology Branch 1
Bureau of Medicine and Surgery
Department of the Navy
Washington, D.C. 20390

35. Chief of Naval Personnel, Pers-11b 1
Department of the Navy
Washington, D.C. 20370

36. U.S. Army Air Defense 1
Human Research Unit
Fort Bliss, Texas 79916

37. Personnel Research Division (Pers-A3) 1
Bureau of Naval Personnel
Dept. of the Navy
Washington, D.C. 20390

38. Doctor Eugene Gloye 1
Office of Naval Research
Pasadena Branch Office
1030 East Green Street
Pasadena, California 91101

39. Motivation and Training Laboratory 1
U.S.A. Manpower Resources Research and Development Center
1300 Wilson Boulevard
Arlington, Virginia 22209

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified	2b. GROUP

3. REPORT TITLE

ESTABLISHING CRITERIA IN PROGRAMMED LEARNING

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Master's Thesis; September 1972

5. AUTHOR(S) *(First name, middle initial, last name)*

Gerald Gerhart Madson
Captain, United States Marine Corps

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1972	60	4

8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

Recent years have seen many attempts to program learning based on many principles ranging from intuition to mathematics. The area of criterion establishment for advancement or conversely additional training at the same or lower levels remains primarily in the realm of intuition. The intuition used is only as good as the experience and background of the program designer however. In the case of a very experienced program designer the criteria may be very efficient. In the case of less experienced program designers the criteria are usually arbitrary. This work describes a method of analyzing learning programs and determining mathematically sound criteria. The mathematical foundation for this analysis is the Markovian learning model as opposed to the linear learning model.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Markovian Learning Theory						
Programmed Learning						
All-or-None Learning Theory						
Criteria						
Advancement Criteria						
Pass Criteria						
Branching Criteria						